Mortality Statistics of the United States Census.

I. INACCURACY.

In no department has the census failed more signally than in the statistics of mortality. The radical difficulty consists in the fact that mortality statistics relate to a *period* of time, whereas the census is taken at a *point* of time. The census methods are fitted only to describe the state of the population on June 1st, not to record the changes in that population during the previous year. Births, deaths and marriages can properly be counted only when they occur, that is, through continuous registration. In the census they are counted by enumeration, that is, by inquiry of the census taker from house to house.

The earliest census in which mortality returns are given is the seventh (1850). The results are the subject of spicy comment in the census volume itself,[1] "Upon the subject of the Deaths no one can be deceived by the figures of the Census, since any attempt to reason from them would demonstrate a degree of vitality and healthfulness in the United States unparalleled in the annals of mankind, would overthrow the best established principles of statisticians, and, in coming down to details as well as in the aggregates, contradict all science and experience." The writer was too sanguine, however, as to the harmlessness of such inaccurate statistics. Experience shows that the census death rates are taken in earnest by hosts of guileless students. Indeed, much worse figures than those of the census of 1850 have been accepted without question. The death rate for whites

[1] Seventh Census, xl.

was given in 1850 as 13.6 per thousand; for slaves, 16.6. But the *Monthly Bulletin* of the Iowa state board of health for June, 1890, presents in perfect seriousness the following death rates: in 1880, 16 per 1,000; in 1883, 3.7; in 1885, 4.5; in 1890, 4.0. The death rate for Dubuque is given as 9.3, for Des Moines, 0.8, and for Council Bluffs, 0.06. These figures are commented upon as follows: "It was not till 1883 that the work of the state board of health began to be realized. The saving of lives, therefore, through the sanitary and protective measures of the state board, is a record to be proud of,"[1] the fact being, of course, that the worthlessness of its statistics is a record to be ashamed of. Inaccurate statistics, unless so labeled, are far worse than none at all. The label should be repeated wherever the figures occur. The tables of 1850, though admittedly incorrect, were thought by Mr. Kennedy to be of value "Should a more critical examination, which time will enable us to exercise, prove the returns of the number of deaths too small, such a result will not affect their value for the purposes of comparison of one portion of the country with another."[2] This optimistic remark is made in face of the fact that the figures show one death to every fifty-one living in Massachusetts, and only one to one hundred in neighboring Vermont!

In every census report the reader is warned against accepting the figures which follow, but the ordinary user of the census volumes will never think of hunting several hundred pages back of a figure to see if it is true or false. Thus, in one place in the Tenth Census,[3] the death rate for the United States is printed 15.09,

[1] Am. Stat. Assn. *Publications*, 2:152 (1890).

[2] Seventh Census. Report of Supt. for Dec. 1, 1852, and Dec. 1, 1851, p. 137.

[3] Tenth Census, 11:3.

although many pages back[1] this figure had been shown to be worthless and 18 is conjectured as the true rate. Why should false figures be given at all? The ordinary reader is certainly excusable if he expects a book of figures to be fact, not fiction. With successive censuses the returns have not improved, but rather degenerated. The official death rate for the United States in 1850 was 14 per 1,000, and in 1890 also 14. But the rate in 1890 was kept up by certain extraneous aids. Instead of trusting solely to enumeration, as in 1850, the censuses of 1880 and 1890 had recourse to the registration records of certain states and cities. Without these records, *i. e.*, within the non-registration area, the official death rate in 1890 was 11 instead of 14. Even another deduction must be made before we can compare this with the accuracy of the enumerators' work in 1850. In 1890, 40,000 deaths were added to the enumerators' record from the supplementary returns of physicians.[2] Without these deaths the official death rate would have been only 10. The character of the record for each census since vital statistics were first gathered is exhibited in the following table:

	Official Death Rate in U. S.	In the Non-Registration Area.	In ditto, but excluding Physicians' Returns.
1850	13.9	13.9	13.9
1860	12.5	12.5	12.5
1870	12.8	12.8	12.8
1880	15.1	13.4	12.0
1890	14.0	10.9	9.9

[1] *Idem.*, xix.

[2] This plan of soliciting corrections from physicians originated in the tenth census. Its extent, in both censuses, is indicated by the following table:

	1880	1890
Letters sent	70,000	75,000
Replies received	26,000	13,000
Deaths added	61,000	40,000

From the last column we see the steady deterioration in the accuracy of the enumerators' work, at least as applied to mortality statistics. Even the inclusion of the supplementary returns from physicians (middle column) did not keep up the last two censuses to the standard of 1850, while the substitution of registration for enumeration in a limited area barely brought them above that standard.

Of course it must not be forgotten that a death rate is the ratio between two numbers, the deaths and the population. An inaccuracy in the population would affect the death rate as truly as an inaccuracy in the number of deaths. The reason why the enumerated death rate in 1870 (3rd column) appears to better advantage than that of 1880 is, perhaps, that the count of population was worse in the former year, not that the count of deaths was better.

Another hindrance exists to making our American statistics comparable with European. The population number used as divisor ought to be, and in Europe is, a mean of the population through the year during which the deaths occur. The population at the middle of the year is usually taken. In the United States, however, that at the end of the year is taken. This population is too large and, in consequence, the resulting death rate is too small. The deficiency from this source amounts to from 1 to 2 per cent.

The above figures show the utter hopelessness of arriving at the general death rate in the United States by a method in which enumeration forms a part. Such a method has been discredited in other countries. The census of 1851 for Ireland attempted to give the

death rate for the preceding decennium by enumerators' inquiry, with the following result:[1]

1842	8.5	1847	34.8
1843	9.0	1848	29.7
1844	9.8	1849	35.1
1845	11.6	1850	24.5
1846	16.8	1851 (3 mos.)	7.0

As this table shows, the more remote the time, the more defective the record. Every event, even a death, is forgotten. Deaths of infants and females especially are overlooked after a certain time. Besides this, many families, especially those consisting of but one or two persons entirely disappear (either by death or emigration) leaving no one to tell the tale.

The same gradation is seen in our census. The deaths reported are less and less, the earlier in the year they come.[2]

If any further evidence of the dismal inaccuracy of the census figures for deaths were needed, it is to be found in a comparison of the results of registration and enumeration. In 1860 the enumerators found 21,300 deaths in Massachusetts; the registration showed 24,100.[3] The deficiency of enumeration compared with registration is here 12 per cent and the registration was doubtless itself defective. In 1870, Mr. E. B. Elliott, the government actuary, in computing his life tables, assumed from the registration of Massachusetts and England that the death rates were deficient in the census returns by 41 per cent.[4] In 1880, the enumerators' returns for Massachusetts (exclusive of Boston) were 27 per cent. short; for New Jersey 37 per cent. Even

[1] U. S. Eighth Census, Mortality and Miscellaneous Statistics, xxiv.
[2] Tenth Census II: xix.
[3] Eighth Census, Mortality and Miscellaneous Statistics, 213.
[4] Ninth Census, Vital Statistics, x.

after the enumerators' returns were supplemented by those from physicians, deficiencies still remained of 13 per cent and 20 per cent respectively.[1] A deficiency of 30 per cent as a maximum was allowed for the whole country. The final conclusion was that "It seems safe to assume that the death rate was not less than 17 nor more than 19 per 1,000 of living population."[2] Consequently the general rate was assumed as about 18. Comparing the registration area with the remainder of the country (where enumerators' and physicians' returns were employed) we find for the last two censuses:[3]

	1880	1890
Registration area	22.2	20.8
Non-registration area	13.4	10.9

We see here that registration gathers nearly twice as many deaths as the other method. It is only fair to add, however, that cities made up a larger proportion of the registration than the non-registration area, especially in 1880, and it is well known that urban mortality is high. But the death rate in the rural part alone of the registration states in 1890 was given as 15.7, and this is too small.[4] Infants' deaths especially escape enumeration (We shall see later that to a considerable degree they even escape registration). For 1890 the deaths under one year per 1,000 infants living under one year were reported as follows:[5]

Registration area	260
Non-registration area	76

In the eleventh census the death rates for separate states were not computed because misleading. In the twelfth

[1] Tenth Census, II: xix.

[2] *Ibid.*

[3] Eleventh Census, Mortality and Vital Statistics, I: 11. These figures differ from those given in Compendium, 2: 4. [For at least a partial explanation see above, p. 112,f.—W. F. W.

[4] Eleventh Census, Vital Statistics, I: 11.

[5] Eleventh Census, Vital Statistics, I: 22.

census we hope no enumeration of deaths whatever will be attempted. Each census apologizes for inserting returns admittedly 20 per cent to 50 per cent short, on the ground that they are of use for the study of diseases which caused death, etc.; in particular for determining the comparative number of different diseases per 1,000 deaths from all causes. When this plea was first made (1850), there were no registration records, and the application of the death returns to such comparisons was probably a true utilization of an otherwise waste product. But in the eleventh census even this excuse for enumeration was of doubtful validity. Much better results were obtainable from registration methods. It is true that defective returns would give correctly the relative number of deaths from different diseases, if the percentage of deficiencies for each disease were uniform. But this is not the case. Each disease varies in frequency with the time of life. Croup predominates in childhood, cancer in advanced life, the diseases of child-birth between twenty and forty. Since the deficiencies of death returns vary with age, the resulting comparisons for diseases must necessarily be inaccurate. Moreover, as stated in the last census, "Such rates [the relative number of different diseases per 1000 deaths from all causes], as is well known to statisticians, have little positive value as compared with the ratio of deaths to population."[1] An epidemic of one disease causes the numbers of all other diseases per 1000 deaths to decrease, without any actual decrease of their frequency.

It must not, of course, be assumed that registration, in and of itself, is necessarily and always accurate. Much depends on the character of the registration laws, still more on the penalties for their violation, and most

[1] Eleventh Census, Vital Statistics, I: 1.

of all on the character of the officials who administer them. Through the kindness of various persons, chiefly secretaries of boards of health, the following rough estimates have been secured of the possible deficiencies in the death-returns of the registration states at the present time.

	Correct within
Maine	3 per cent
New Hampshire	1 "
Vermont	3 "
Massachusetts	1 "
Rhode Island	1 "
Connecticut	1 "
New York state[1]	5 "
New York city	1 "
New Jersey	1 "
Delaware	10 "
District of Columbia	1 "
Michigan	5 "

A few of those credited with a one per cent accuracy are probably deserving of a still closer estimate. For the registration states as a whole, it is probably safe to say that the deaths are reported correctly within two per cent or less.[2] The death *rate*, involving as it does the number of population, is probably correct within two or three per cent for the census year and within two to six per cent for intercensal years. The Census Office did not include within its "registration states" all states professing that name. Alabama claimed to have complete registration in 1890, but the registration death rate was only 13.8, as against a mere enumeration rate of 14.2 in 1880![3]

[1] Exclusive of New York city, Brooklyn, Yonkers, Albany and Buffalo.

[2] Similar estimates were secured for births and marriages. Birth returns are supposed to be deficient by 1 per cent in Rhode Island, 2 per cent in Vermont and Massachusetts, 3 per cent in Maine, and 5 to 25 per cent in the other states. Marriages range from 1 to 30 per cent.

[3] See W. A. King, Vital Statistics of the Census. In Am. Stat. Assn. *Publications*, 5: 214 (1897).

In one city in New York the death rate appeared as 13 per 1000. But on inquiry by Mr. Willcox it was found that a quarter or a third of the deaths escaped record, because of the failure to require burial permits before interment.[1] In a population sparser than forty to the square mile, complete registration is regarded as out of the question.[2] In any community the first years of registration do not give trustworthy results. Massachusetts has had registration since 1842, but the interesting "Vital Statistics of Massachusetts, 1856–95," begins with 1856 because the earlier records were thought to be of doubtful value. For the forty years, 1856–95, the records are regarded as "approximately correct."[3] Accuracy is everywhere a matter of growth. In England Farr estimated the probable deficiency in the registration of births for 1861–70 as one-sixth of one per cent, but for the period 1841–50 as three times as much.[4] The infant death rate is unreliable even under the most perfect system of registration. This is due to deficiency in the number of infants recorded in the population quite as truly as to deficiency in the number of deaths recorded. The deficiency in enumerated living infants is, in turn, partly due to the fact that infants 9 to 12 months old are returned as "one year old." This is true in all countries.[5]

One method of measuring such inaccuracy can be briefly described as follows: Take any particular year and make abstraction of migration; then the number born in that year, less those born and dying in it, is the

[1] *Idem.*, 5: 202.

[2] See C. L. Wilbur, *Idem.*, 191.

[3] Vital Statistics of Massachusetts, 1856–95, 714. In Mass. State Board of Health Reports, 1896.

[4] Farr, Vital Statistics, 523.

[5] Eleventh Census, Vital Statistics, I: 21.

number under one year old living at the end of the year. In Massachusetts, however, while the number of infants under one year old, thus computed, was found in 1890 to be 53,000, the number actually enumerated in the census was only 43,000![1] Adding to the 53,000 a supposed correction of 2 per cent. for deficiencies in the registration of births, we find the discrepancy still greater, or 26 per cent. Similar calculations in 1880 gave a discrepancy of only 5 per cent. Taking an interval of five years, in place of one year, we find in 1890 a deficiency of 14 per cent.; in 1880, of 3 per cent. These figures show the lamentable inaccuracy of the eleventh census, as well as the fact that even the tenth left much to be desired. This is seen by comparing the English deficiencies with those for America in 1880. In any population (especially in an increasing one, as in the United States) the numbers living between the ages 0–1 should exceed those 1–2; those 1–2 should exceed those 2–3; etc. But enumerators' statistics always show the opposite. Hence we have a rough measure of deficiency by subtracting the numbers returned as 0–1 from those returned as 1–2; those 1–2 from those 2–3; etc., and expressing the differences as percentages of the smaller numbers. The results are as follows:[2]

	Deficiencies (per cent).	
	U. S.	Eng. and Wales
Age 0–1 compared with 1–2	13	9
" 1–2 " " 2–3	13	3
" 2–3 " " 3–4	3	2
" 3–4 " " 4–5	1.4	0.7

Corresponding discrepancies exist for age divisions by months, especially in the eleventh census. The number returned as under one month was less than a seventh of

[1] Eleventh Census, Vital Statistics, I: 490.

[2] Tenth Census, 12: clvi.

those returned as from one to two months old.[1] The deficiencies vary for different sections and classes. Thus, for age 1–2 as compared with 2–3 the deficiencies were: whites, 13 per cent; colored, 18 per cent; native whites, 12½ per cent; foreign born, 35 per cent. For New York it was 16 per cent; Louisiana, 30 per cent; New Mexico, 46 per cent.[2] That is, enumeration of population encounters the worst obstacles among negroes and foreign born and in the sparsely settled west. Although the tenth census points out inaccuracies in infant mortality, it does not hesitate to make international comparisons. In the ninth census the government actuary, Mr. E. B. Elliott, gives a formula for correcting the irregularities in the returns for infants.[3]

II. LIMIT OF ACCURACY.

We must not make the mistake of classifying statistics into accurate and valuable, on the one hand, and inaccurate and valueless on the other. No measurements are truly accurate. The important thing is to know the degree of inaccuracy, or at least the direction of inaccuracy. To be sure that a figure is too high or too low is sometimes enough to make it of value. It becomes an upper or lower limit. If the negro death rate is known to be sadly deficient while the white death rate is nearly correct, the former will still be useful if, as given, it is greater than the latter; for then we are doubly sure that the true negro mortality exceeds the white. Even the early Greek computation that the sun was at least 19 times as far away as the moon was a real addition to knowledge, though the true number is about 400.

[1] Eleventh Census, Compendium, 3: 198.
[2] Tenth Census, 12: clvi.
[3] Ninth Census. Vital Statistics, 1: 517,ff.

When both an upper and a lower limit can be given we are much better off. The estimate that the general death rate in the United States was, in 1880, between 17 and 19 per 1000,[1] placed it above that of Norway and below that of Italy. In almost every case, some estimate can be given of the amount and direction of possible error. The bane of ordinary "Statistics" is the omission of this estimate. Where the errors are large, they should be given in footnotes accompanying the table. Where they are less than 5 per cent, they can be sufficiently indicated by printing only the significant figures (and ciphers as needed).

The importance of stopping at the last significant figure is not yet recognized among statisticians. The very phrase is strange to most of them. Its meaning will appear from a few examples. The official population of the United States on June 1, 1890, was 62,622,250. A very intelligent gentleman expressed surprise when told that this figure was not correct to the units place. If it were correct, it would be the most accurate measurement ever accomplished. No chemist's balances, no astronomer's calculations have ever yet reached a precision of one part in sixty million. The most accurate physical measurements are correct only to one part in a hundred thousand to a million, *i. e.*, require only six significant figures, and this extreme degree of accuracy is quite out of the question in statistics. Not more than three significant figures could possibly be claimed for the population count (62,600,000) and, probably, two are amply sufficient[2] (63,000,000). Few statistics are reliable beyond two and few are useful beyond three significant figures. We should prac-

[1] Tenth Census, II : xix.

[2] Cf. Am. Stat. Assn. *Publications*, 4 : 99 (1894–95).

tically confine ourselves, therefore, to a choice between two and three. Where the second figure is certain and the third uncertain within only a few units (say three), three places should be used; in case of greater uncertainty, two. If the uncertainty amounts to say 5 per cent or more, the method of a separate upper and lower limit should be employed.

Few people realize the need of these rules in statistics, the extent to which useless figures are tacked on, and the false appearance of precision given to the most hypothetical calculations. The gold in the United States on July 1, 1896, was officially estimated at $599,597,964. This result was found by guessing at the gold in 1873 as twenty millions and applying the subsequent reported imports, exports, recoinage and consumption. The final result is worthless beyond one significant figure.[1] Such official computations appear to be based on the theory that a sum is as accurate as the most accurate of its parts. The truth is that a sum cannot be more accurate than the *least* accurate of its parts.

Another important principle is that to multiply or divide a figure by a constant cannot improve its degree of accuracy. A writer on the Spanish dollar[2], translating from one unit to another, brings out of an original four place figure the result that the Castilian mark contained 230.0465 metric grams. As the coin in question related to the period of the discovery of America, and as even our modern mints can scarcely attain such a degree of accuracy, the non-significant character of the last three figures is apparent.

Index nnmbers are often expressed with four or five digits when not more than two have meaning.[3]

[1] Yale Review, 7: 105 (1898).
[2] Am. Hist. R., July, 1898, p. 607.
[3] Cf. A. L. Bowley, *Economic Journal*, 5: 374, (1895.)

In the special field of mortality statistics, examples of false accuracy abound. In an article on Vital Statistics of an Apache Indian Community[1] we find a death rate of 142.58. This is based on a population of 371 and on 53 deaths! The colored death rate in Rochester is elaborately tabulated in the eleventh census as follows:[2]

All ages	6.92
Under five	21.28
Five years and over	5.65

These figures are based on a colored population of 578 and on 4 deaths! This fact is stated below the table, but only in the reading matter and, apparently, because the results are so peculiar. Of course such statistics have no significant figures at all. They should be omitted. The same is true, as already shown, of the general death rate in the United States (15.09).

Intercensal rates are less accurate than those of the census year, for the population must be guessed at ("estimated") on the basis of the previous rate of increase. In England, in the census year 1891, the estimated population was found to be 2½ per cent. in excess.[3] Such an error invalidates even the third significant figure (whether of population or of death rate based thereon). In the United States the probable errors in such estimates are much greater.

Statisticians have juggled with death rates by assuming a great increase of population. Thus in St. Paul,

[1] Am. Stat. Assn. *Publications*, 3 : 427 (1893).

[2] Eleventh Census, Vital Statistics, 2 : 44.

[3] The English have it in their power to improve their estimates by adopting a different method, viz., to add to the last census number the intervening births and immigrants, and subtract the deaths and emigrants. By this method, Dr. Longstaff calculated the population in 1891 correctly to *four* significant figures. Unfortunately this method can not be used in the United States. See Cannan, "Demographic Statistics of the United Kingdom; their want of Correlation and other Defects." In Royal Statistical Society, *Journal*, 61 : 50 (1898).

1896, by estimating the population as 216,000, as against 155,000 in 1894, a death rate of 6.7 was presented as against 10.1 in 1894.[1]

III. DEATH RATES.

Though the desired statements as to accuracy are lacking, there can be no doubt that, in general, the registration portions of the census volumes on vital statistics are of much value. The last two censuses are the only ones which have made any considerable use of state and municipal records. Indeed it is only in recent years that these records have been available in sufficient quantity and quality. In the tenth census two volumes were devoted to vital statistics; in the eleventh, four, (of about 1100 pages each). Vol. I treats of general analysis and rate tables, vol. II of the vital statistics of cities of 100,000 and upward, and vols. III and IV of the detailed statistics of deaths by locality, etc., the raw material, as it were, of the previous volumes. Besides these there are three special volumes on the cities of New York and Brooklyn, Boston and Philadelphia, and Washington and Baltimore, which are, however, practically included in one of the above four. They cover about 1000 pages in all. A large portion of these seven volumes is devoted to registration matter, and this is not simply transcribed from the state and municipal reports. It often happened that many items returned on the original cards were not entered in the printed or even manuscript reports of the separate states or cities. In such cases the census directed a thorough examination of the original cards. Thus, the census report, though derived from

[1] Am. Stat. Assn. *Publications*, 5 : 362 (1897).

anterior state and municipal sources, contains much which is not to be found there or elsewhere.

In the tenth census the registration states were Massachusetts, New Jersey and the District of Columbia. The "registration area" consisted of these states and 19 cities in other states. In the eleventh census the "registration area" consisted of the registration states, New Hampshire, Vermont, Massachusetts, Rhode Island, Connecticut, New York, New Jersey, Delaware, and the District of Columbia, and of 83 cities in the non-registration states with "fairly satisfactory systems of registration." The total number of cities included was 271, of which 28 had a population of 100,000 or more, viz.: New York, Chicago, Philadelphia, Brooklyn, St. Louis, Boston, Baltimore, San Francisco, Cincinnati, Cleveland, Buffalo, New Orleans, Pittsburgh, Washington, Detroit, Milwaukee, Newark, Minneapolis, Jersey city, Louisville, Omaha, Rochester, St. Paul, Kansas city (Missouri), Providence, Denver, Indianapolis, Allegheny. In the twelfth census the number of registration states will be increased by Maine and Michigan and possibly others, the most likely being: Maryland, Indiana, Iowa, and Pennsylvania.

The death rates (including still births) in the registration states were as follows:[1]

Vermont	16.3
Delaware	18.4
New Hampshire	18.8
Connecticut	19.4
Massachusetts	20.2
New York	20.5
New Jersey	21.0
Rhode Island	21.9
District of Columbia	25.9
Total Registration States	20.4
Total Registration Area	20.8

[1] Eleventh Census, Vital Statistics, I: 10.

The registration area contained a population of 21,000,000.

In the registration states, the urban death rate was given as 23.5 (supposed to be nearly accurate) and the rural 15.7 (supposed to be too low, the actual rate being, say, 16).[1] The rate varies directly with the size of the city as the following table shows:[2]

Cities.	Males.	Females.
Under 10,000	19.2	16.9
10,000– 15,000	19.3	17.2
15,000– 25,000	21.2	18,7
25,000– 50,000	23.2	20.4
50,000–100,000	23.7	20.9
100,000 and over	25.1	21.5

Since size corresponds usually to density, this table agrees with the observations of Farr in England.[3]

Number of Districts.	Persons per square mile.	Death rate.
54	170	15 to 17
349	200	18 to 20
142	450	21 to 23
56	2,200	24 to 26
16	6,900	27 to 30
1	12,000	32
1	66,000	37

The statistics for cities[4] over 100,000 show, in 1890, an average mortality (exclusive of still births) of 21.6. The highest record is for Newark (27.4), with New York second (26.5), followed by Jersey city and Brooklyn. The lowest is Omaha (9.4), for which the record is probably defective. The death rates in the 28 largest cities in United States, England and Germany, in 1890 were:[5]

United States	21.6
England	21.4
Germany	23.1

[1] Eleventh Census, Vital Statistics, I: 11.
[2] Eleventh Census, Vital Statistics, I: 220.
[3] Farr, Vital Statistics, 126.
[4] Eleventh Census, Vital Statistics, II: 2.
[5] *Ibid.*

Remembering that the American rate is too small (owing to the use of the population at the end of the year instead of the middle, and to some deficiency in the record), it may fairly be said that the death rate for large cities is the same in this country as in England and Central Europe. The death rate of Paris in 1890 was 24.5, and of Vienna, 24.6. In Russia and Italy larger rates are found: St. Petersburg, 28.4; Rome, 33.4; Moscow, 40.3. The rate for Havana is said to be from 70 to 140!

The true rate for the United States as a whole is supposed to be about 18. But this appears to be a mere repetition of the surmise of 1880.[1] It is not likely that the rate in 1890 should be the same as in 1880. In every country in Europe, except Norway, the rate in 1890 was less than in 1880. On the other hand, in the registration states common to the two censuses we find an increase:[2]

	Massachusetts.		New Jersey.	
	Males.	Females.	Males.	Females.
1880	19.1	18.1	17.0	15.7
1890	20.9	19.4	22.3	19.7

This increase is probably due to an increasing proportion of urban population (and in part to greater accuracy in registration).

As is seen above, the female death rate is less than the male. This was found in 1890 to be the case in every registration state except Vermont. For the registration states taken together, the results are: males, 21.5; females, 19.3. The excess of male mortality is well known in Europe. The higher death rate for males compensates for the higher birth rate. In a stationary population, according to the English laws of mortality and natality, the sexes would be evenly distributed. There would be 512 male births for every 488 female

[1] Eleventh Census, Vital Statistics, I: 11.

[2] Eleventh Census, Vital Statistics, I: 13.

and these would sustain a population of 20,430 males and 20,430 females.[1]

The poor have a much heavier death rate than the rich. Châteauneuf, comparing the rich with the poor quarters of Paris, found the following mortality:[2]

Age.	Rich.	Poor.
30–35	0.85	1.4
40–45	0.85	1.9
50–55	1.8	2.6
70–75	6.8	14.1

Similar contrasts result from the returns of industrial insurance (exclusively among the poor) compared with ordinary insurance. The following table of mortality is given by the Vice-President of the Metropolitan Life Insurance Company (industrial):[3]

Age.	Ordinary Insurance. (English Experience.)	Industrial. (Metropolitan.)
20	7.3	10.5
25	7.8	14.1
35	9.3	17.2
55	21.7	35.2
70	64.9	91.0

Like conclusions are reached by inspecting the death rate chart for separate wards of New York.[4] That the upper classes live longer than the average is also pointed out in the *Journal of the Royal Statistical Society*.[5]

The mortality among the colored population is very heavy. The eleventh census showed for the registration area the following results for different ages:[6]

	Under 5.	5–15	15–45	45–65	65 and over
White	64	5.4	9.3	22	78
Colored	124	10.4	17.0	33	106

[1] Farr, Vital Statistics, 485.

[2] A. Quetelet, Physique sociale, I: 361.

[3] The Charities Review, March, 1898, p. 33.

[4] Eleventh Census, Vital Statistics of New York and Brooklyn, 164, 250.

[5] 50: 649.

[6] Eleventh Census, Vital Statistics, I: 19.

In ten southern cities the white death rate is 20.1 and the colored, 32.6.[1] We must remember, however, that the deficiency in the population returns is greater among the colored than among the whites, especially in infancy. Yet the general facts as to the relative mortality between the two races in the United States cannot be doubted, and are abundantly confirmed by Hoffman's monograph[2] and by the experience of Insurance Companies.[3] The death rates by nativity in registration states in 1890 were as follows:[4]

Native whites of native parents	17.0
Native whites of foreign parents	24.4
Foreign born whites	19.9
Colored	19.6

These figures, however, are not comparable owing to differences in age distribution.

Mortality varies with age. It is greatest in infancy and old age and least in childhood. In the registration area we find the death rates per 1000 to be as follows:[5]

Under 1.	Under 5.	5 to 15.	15 to 45.	45 to 65.	65 and over.
206	67	5.6	9.7	22	79

Infant mortality is the subject of special study. It also presents special difficulties. Not only are the returns of infant deaths deficient, but the returns of infant population, on which death rates must be computed, are also deficient. On account of such deficiencies, special modes of reckoning infant mortality have been adopted. For instance, the deaths of infants under one year old occurring during a year are compared, not

[1] F. L. Hoffman, Race Traits and Tendencies of the American Negro, 39.

[2] *Op cit., passim.*

[3] Am. Stat. Assn., *Publications*, 3 : 350 (1893).

[4] Eleventh Census, Compendium, II : 4.

[5] Eleventh Census, Vital Statistics, I : 15.

with the population living under one year old, but with the births registered during the year.

In the United States, however, this method is not regarded as accurate because of the inaccuracy in the registration of births.[1] Another method, not of great value, but adopted in the census, is to compare the deaths of infants under one year old with the total deaths.[2]

Still births are often included in the birth and death rates. They are sufficiently numerous to make a difference in the second significant figure. In reckoning the infant death rate especially, still births make an enormous difference. Thus, in the registration area for 1890 we find the number of deaths under one year per 1000 births is:[3]

Including still births	221
Excluding still births	183

A very little consideration will show that still births should not be included in births and deaths. They represent antenatal deaths, and should be studied in connection with antenatal population. Ideal vital statistics should begin with conception, not with birth. In such statistics room would be found for miscarriages and still births. Since, practically, such statistics are out of the question, the proper mode of treating still births is as a separate category.

IV. "CORRECTED" DEATH RATES.

Few cautions are more necessary to the student of vital statistics than to beware of the disturbances of age distribution. The foreign born population of the

[1] Eleventh Census, Vital Statistics, II: 30.

[2] Eleventh Census, Vital Statistics, I: 22.

[3] Eleventh Census, Vital Statistics, I: 21.

United States have had to suffer unjustly in statistical comparisons from disregard of this caution. Their criminality, for instance, appears high simply because of the small number of children and old men, and the large number living at the intervening or "criminal" ages. Similarly the longevity of men of note (such as Gladstone and Bismarck) is often interpreted to mean that distinction and heavy responsibilities lengthen life. But men of note are *ipso facto* old to start with, while obscure persons are largely made up of the comparatively young.

Sometimes a rough correction for age distribution is given by excluding certain ages from the comparisons altogether. Thus, Thomas A. Welton, in treating of "Local Death Rates in England and Wales, 1881–1890," attempts to show that "between the ages 5–45 and particularly between the ages 15–35, such an impression is made upon death rates by migration and by casual events that, for the purpose of measuring the strength of the force operating for and against life, the facts for those ages should be left out of account, and attention should be concentrated on statistics as to ages 0–5 and 45 upwards."[1] Similarly the birth rate is often reckoned in relation to the female population of "child-bearing age," *i. e.*, 15–50,[2] and the mortality among different occupations compared by confining attention to ages over fifteen. But such crude methods are often inadequate. Thus, in 1890 the death rate among "professional" men over fifteen was 15.7 and among the "manufacturing" class 13.0, indicating, apparently, a greater vitality

[1] Royal Statistical Society *Journal*, 60:33. Summarized in Am. Stat. Assn. *Publications*, 5 : 363, (1897).

[2] *E.g.*, Eleventh Census, Vital Statistics, II : 32.

among the latter. But a minuter comparison, for different age groups, shows the opposite:[1]

	15–25.	25–45.	45–65.	65 and over
Professional	5.0	8.5	19.1	79
Manufacturing	5.0	9.2	20.1	78

Great masses of the mortality statistics in relation to occupation given in the census volumes are, therefore, of no value. A still more striking example may be drawn from the mortality of different conjugal conditions. In 1890, in 22 cities, the death rate among males over 15 was: single, 13.7; married, 19.4,—an apparent argument for the single state. But a little consideration shows that the married are generally older than the single. The misleading character of the preceding comparison can be exposed by using only two age groups:[2]

	15–45.	45 and over.
Single	11.7	52
Married	11.4	34

In the same way it is found that though the frequency of marriages is greater among spinsters than widows yet for each age group taken separately, the opposite is true.[3] Even when the correction for age distribution does not reverse the conclusion, it is important in affecting its force. Thus, the average mortality among whites in New York for six years was 26.7, and among colored, 28.8,[4] which contrast is much less than usual (*e. g.*, for the registration area, white, 20.2; colored, 32.4). The reason for the small difference in the rates for the metropolitan area is the presence of so many adult colored waiters and servants, and the comparative absence of infants. Again, the urban death rate, though usually greater than the rural, would be still greater if the two

[1] Eleventh Census, Vital Statistics, I: 65.
[2] Eleventh Census, Vital Statistics, II: 52–54.
[3] R. Mayo-Smith, Statistics and Sociology, 109,f.
[4] Eleventh Census, Vital Statistics, I: 30.

populations were similarly constituted as to age. Thus, for England:[1]

	Death Rate.	
	Uncorrected.	Corrected.
City	24	24
Country	19	16

Where the direction of the error arising out of unlike age distributions is subject to conjecture, the comparison is rendered worthless. This is true of the following comparisons, giving the death rates (in the rural parts of the registration states) according to birthplace of mother.[2]

Mother born in	Death rate.
United States	11.7
England and Wales	14.4
Germany	12.9
Italy	10.2

A "corrected" or "refined" death rate is based on an assumed standard age distribution. It is the death rate which would exist in the given population if its age distribution conformed to that standard. Suppose, for instance, we wish to compare the mortality of French and English children. An actual population group of 10,000 of each is found. Let us assume that in the English group 400 die per annum and in the French only 235. The English crude death rate is therefore 40 per 1000 and the French 23.5. But the two groups are very differently constituted, viz.:

	English.		French.	
Ages.	Population.	Deaths.	Population.	Deaths.
0 to 5	5000	350	1000	100
5 to 10	5000	50	9000	135
	10,000	400	10,000	235

Whereas the English children are equally divided be-

[1] Mayo-Smith, Statistics and Sociology, 151, f.

[2] Eleventh Census, Vital Statistics, I: 36.

tween the two age groups, the French have nine times as many in the upper age group as in the lower. Consequently among the English the younger deaths predominate; among the French the older. For the age group 0–5 the English death rate is 350 per 5000, or 70 per 1000; for the French the corresponding rate is 100 per 1000. For the ages 5–10 the English rate is 50 per 5000, or 10 per 1000; the French, 135 per 9000, or 15 per 1000. That is, in both cases, the French rates are *higher*, viz:

	Death Rates per 1000.	
Ages.	English.	French.
0 to 5	70	100
5 to 10	10	15

The crude rates for the two nationalities were thus misleading. We are to seek a refined rate on the basis of *similar age distribution.* Let us, for instance, take the English age distribution as standard and find out what the French death rate *would* be with a similar distribution (half in each age class):

	French.				
Ages.	Assumed Age Distribution.		Known Death Rates.		Deaths.
0 to 5	5000	×	100 per 1000	=	500
5 to 10	5000	×	15 per 1000	=	75
	10,000				575

There would have been 500 deaths in the first class and 75 in the second. The total is 575 for an assumed population of 10,000, or 57.5 per 1000. This is the refined French death rate made comparable with the 40 per 1000 for the English, and is, as it obviously ought to be, the greater. We may, equally well, take the French age distribution as standard and assimilate the English to it.

This would give for England a refined death rate of 16 per 1000 to be compared with the French 23.5 per 1000; or, we may assimilate both populations to some third standard. Let us assume, as such a standard, two age classes in the ratio of 2 to 3, or 4000 to 6000:

ENGLISH.

Ages.	Standard Distribution.		Known Death Rate.		Deaths.
0 to 5	4000	×	70 per 1000	=	280
5 to 10	6000	×	10 per 1000	=	60
	10,000				340

FRENCH.

Ages.	Standard Distribution.		Known Death Rate.		Deaths.
0 to 5	4000	×	100 per 1000	=	400
5 to 10	6000	×	15 per 1000	=	90
	10,000				490

This gives an English rate of 34 per 1000, and a French rate of 49 per 1000.

For purposes of international comparison it matters little which standard is taken. In the first case (English standard) the French death rate exceeds the English by 44 per cent., in the second (French standard) by 47 per cent., and in the third (arbitrary standard) by 44 per cent. In actual practice the standard chosen for age distribution is of even less consequence than in this imaginary case. For the purpose of comparing the mortality in the different nations of Europe, the age distribution of Sweden has been adopted as standard. The average age constitution of Europe has also been taken. The results of the two standards usually agree within one-half of one per cent.[1] Dr. Josef von Körösi, of

[1] Am. Stat. Assn. *Publications*, 3: 457 (1893). The age groups should be so contrived as to distribute properly the ages ending in 0 or 5. A great number of persons return their ages only in round numbers. Compare above p. 57, f.

Budapest, has emphasized, more than any other writer, the importance of correcting the crude death rate. He had, however, been anticipated by Dr. Ogle in England and Dr. Koch in Germany (Hamburg), and still earlier by Dr. Jarvis, of Dorchester, Mass. Ogle proposed 12 age groups, viz.: each year up to 5, 5 to 10, then each ten years to 60, and lastly 60 and over. Körösi showed that four age groups were enough, viz.: 0–1, 1–20, 20–50, 50 and over. The results of the two calculations differ usually about one per cent. The International Institute of Statistics, meeting at Berne in 1895, recommended five age groups and the Swedish population as the standard. Thus:

Age Groups.	Swedish Population in each 100.
0–1	2.6
1–20	39.8
20–40	27.0
40–60	19.2
60–	11.5

The census has made little use of refined death rates. In the eleventh census the subject was referred to with disfavor: "In the large cities of the United States a correction of gross death rates for peculiarities of race distribution of the population of each city would be much more important than the above correction for age distribution, but it seems hardly worth while to try to indicate the relative healthfulness of different cities by rates for the total population only, and the phrase 'corrected rates' must always be more or less misleading;"[1]

It is scarcely an argument against eliminating one source of error, that another source is more important. Nor is it certain that race distribution is more important than age distribution. It may seem so because, in our cities, different races have different age distributions.

[1] Eleventh Census, Vital Statistics, II: 30.

The true view would seem to be that age, sex and race distributions should all be considered. Different problems require different elements to be eliminated. A death rate corrected for age distribution is only one of many death rates of importance to the statistician. In pressing the claims of a refined rate it is not pretended that even crude rates have no value. For instance, to estimate the rate of increase of a population the birth rate is to be compared with the crude death rate.

The importance of a refined rate is especially great in America, for the age distributions of our various population groups are widely diverse. Thus, we find:[1]

	Death Rate.	
	Gross.	Corrected.
Denver	24.2	31.0
San Francisco	23.6	29.9
Omaha	10.0	14.7

The results of correction may, of course, be in either direction. Thus, Chas. E. Burnap, writing on "Mortality in Twenty-Three Massachusetts Cities,"[2] finds:

	Death Rate.	
	Recorded.	Corrected.
Boston	24.6	26.0
Worcester	19.4	19.4
Salem	20.6	19.8

V. AVERAGE LIFE-TIMES.

All the systems of weighting hitherto described, by which a corrected death rate is obtained as an average of

[1] Eleventh Census, Vital Statistics, II: 30. The correction is according to the Berne recommendations.

[2] Am. Stat. Assn. *Publications*, 5: 82 (1896). The method employed by Burnap is a variation of the method described above. He first compares, say, Boston with Massachusetts in general by taking the Boston age distribution as standard. Next Worcester is compared with Massachusetts by the Worcester standard, and Salem by the Salem standard. Having thus effected comparisons between each city and Massachusetts, these cities are compared with each other.

the death rates at separate ages, are quite empirical and arbitrary. But the rational student will crave a more rational system. The theoretically true corrected death rate is based on an age distribution, given in the nature of things, not the Swedish, not the average of Europe, but that of an ideal stationary population.[1] This true corrected death rate has the important property that it is the reciprocal of the average duration of life. Death rates and length of life are thus opposite facets of the same fact. Mortality measures vitality.

In England, statistics for a period of 35 years showed an average crude death rate of 22.4 per 1000; the true corrected rate was 24.5, *i. e.*, $\frac{24.5}{1000}$.* This shows an average life time of $\frac{1000}{24.5}$ or about 41 years.

The connection between death rate and life time is so important that the following outline demonstration of their relation is given:

Suppose,

100 persons born daily, and that
99 survive one day,
85 survive one year,
70 survive ten years,
50 survive forty years,

and so on, until all are dead. Such a table is called a table of survivors, or a life table.

Suppose, for convenience, that each day's births (and deaths too) occur at the beginning of that day. It is evident that on any one day the population must consist, and consist solely, of

[1] Cf. Farr, Vital Statistics, 121.

* *Ibid.*, 124.

the 100 persons just born,
the 99 persons survived from yesterday's 100 births,
the 85 persons survived from the 100 births one year ago to-day,
the 70 persons survived from the 100 births ten years ago to-day,
the 50 persons survived from the 100 births forty years ago to-day.

This is an age distribution table. Its necessary numerical identity with the survivorship table is apparent. In actual populations, however, the two tables are different except, perhaps, in France, which has a nearly stationary population. We suppose each table to be completely filled out, down to successive days, and not with gaps as above. Evidently the sum totals of the two tables are equal. Suppose this total to be 15,000,000. The total of the age distribution table is clearly the total population. This is, therefore, 15,000,000 and remains constant. Hence, the daily deaths must equal the births (100). The death rate each day is therefore $\frac{100}{15000000}$.

Next, we observe that the sum total of the survivorship table is the total number of days lived by the 100 persons born on any one date; for they each enjoy one day's life or 100 days together. Then 99 are left, who all live a second day or 99 days together, and so on, making 15,000,000 days in all. The average life time of the 100 persons is therefore $\frac{15000000}{100}$ days. Thus the reciprocal relation between average life time $\frac{15000000}{100}$ and true corrected death rate, $\frac{100}{15000000}$, is apparent. The relation remains intact if the life time is reduced to years instead of days and the death rate to a yearly instead of a daily rate; for these changes merely introduce 365 as divisor and multiplier.

To correct the death rate, then, in a community of abnormal age constitution, we make it normal by sup-

posing it to be fed by a constant number of births. The death rate which would exist if the population were thus kept stationary is the true corrected death rate. To obtain it in any community we need a table of survivorship.

Since such death rate is interchangeable with the mean duration of life, it will be convenient to pass to the latter conception.

The mean duration of life, in most civilized countries, is between 35 and 50 years (corresponding to a true corrected death rate of between 25 and 20 per mille). Thus:

Norway (1856–65)_____	47 years,	France (1877–81)__	41 years,
England (1881–90)_____	44 years,	Belgium (1841–50)_	37 years,
United States (1830–60)__	42 years,	Prussia (Becker)__	37 years.

The mean life time is on the increase.

England,[1] (1838–54)___	39.9	France,[2] (1856–65)_____	39
" (1871–80)___	41.4	" (1877–81) (males)__	41
" (1881–90)___	43.7		

Life tables for Sweden show an increase of fifteen years since the middle of the last century. The city vitality especially has improved. The mean lifetime in Stockholm, 1755–63, was only 16 years.

In Massachusetts the death rate has remained stationary for four decades:[3]

1856–65	19.4	1876–85	19.4
1866–75	19.6	1886–95	19.6

This, however, is due to an increase in urban population. The death rate in both city and country has declined.[4]

[1] Vital Statistics of Massachusetts, 1856–95, 815. In Mass. State Board of Health Reports, 1896.

[2] Levasseur, Tables of Mortality and Survivorship. In Royal Statistical Society *Journal*, 50: 547 (1887). Translated from *Journal de la Société de Statistique de Paris*, March, 1887.

[3] Vital Statistics of Massachusetts, 1856–95, 743.

[4] *Idem.*, 748, f. The opinion that American vitality is declining appears to be based on a wrong method of estimating mean life time, viz., by taking the average age of the dying. Such methods seem to

The gravitation of population towards cities tends throughout the world towards increasing mortality, while improved sanitation, and medical knowledge and skill tend directly opposite. The magnitude of the latter forces is, therefore, understated in any comparison between old and new mortality which does not distinguish between city and country.

The steady improvement of vitality is also apparent from the following condensed tables of survivorship for the Netherlands, Sweden and France:[1]

NETHERLANDS.

Age.	Baumhauer. 1840–51.	Von Pesch. 1870–80.
0	1000	1000
10	644	654
60	310	357

That is, out of each 1000 persons born, 47 more now survive to be 60 years old, than 30 years ago.[2]

SWEDEN.

Age.	Wargentin. 1757–63.	De Berg. 1861–70.
0	1000	1000
10	611	737
60	293	401

In Sweden, 108 more persons out of each 1000 now survive to 60, than a century ago. The infant mortality was reduced one-half in the same period.[3]

FRANCE.

Age.	St. Maur. Before 1750.	Demonferrard. 1817–32.	Bertillon. 1856–65.
0	1000	1000	1000
10	484	668	681
60	168	365	389

vitiate the conclusions of John Stockton-Hough, Statistics relating to the Births, Deaths, Marriages, in Philadelphia, 1861–72. Social Science Assn. of Phila., 1874, p. 14,ff. For the fallacies of the method, see Farr, Vital Statistics, 475,ff.

[1] Levasseur, Royal Statistical Society, *Journal*, 50 : 558 (1887).

[2] Levasseur, *Ibid.*, 555.

[3] Levasseur, *Ibid.*, 553.

Here, in about 75 years, 197 more persons were preserved to age 60 than previously. In the next 35 years the saving was 24. The rate of improvement is thus decreasing.

The improvement is so marked that even crude death rates are sufficiently exact to show it strongly. In London, in the latter half of the 17th century, the death rate was about 80 per 1000 and in the 18th century it was about 50, while in the middle of the present century it was 24.[1]

In the 18th century the death rate in Boston varied (1725–74) between 33 and 43, averaging 37.[2] To-day it is about 25.

The effect of improvements in medical science and sanitation is well illustrated by the disappearance of sudden jumps in the death rate through epidemics. In London, in the plague years 1593, 1625, 1636, 1665, the death rates per 1000 were 240, 310, 130, 430.[3]

The introduction of a filter, in 1893, into the water supply of Lawrence, Massachusetts, was followed by an immediate reduction in the deaths from typhoid fever.[4]

Years.	Death Rate from Typhoid Fever.
1887	1.2
1888	1.2
1889	1.4
1890	1.3
1891	1.2
1892	1.1
1893	.87
1894	.50
1895	.31
1896	.19

[1] Farr, Vital Statistics, 131.

[2] Bills of Mortality, 1810–49, City of Boston, Reprinted for Registry Department, 1893.

[3] Farr, Vital Statistics, 131.

[4] Mass. State Board of Health Report, 1896, 568.

The introduction of sewerage and house drainage was followed in ten cities by a reduction of 69 per cent in the death rate from typhoid fever.[1]

The general death rate in New York city in 1896, when Colonel Waring had charge of the street cleaning, was 21.5. The lowest annual rate in the preceding decade was 22.8 and the average, 25.2.[2]

It was found in Massachusetts that while infectious diseases had grown less deadly between the period 1875–84 and 1885–94, local diseases had grown more so, leaving the total mortality about the same.[3]

The life time in city is less than in country. Contrast *e. g.*[4]

Surrey, males	44	Massachusetts and New Jersey	45
Liverpool, males	25	New York city	33

Male life time is less than female.[5]

	Male.	Female.
England, (1876-80)	42	45
United States	41.0	42.9
New Jersey, 6 years	40.1	43.6

The mean life time is also called (rather infelicitously) the expectation of life at birth. The expectation of life may be taken at other times than birth. It is then called the mean after-life-time. Just as the mean life-time is the reciprocal of the true corrected death rate, so the mean after-life-time, for any particular age, is the reciprocal of the true corrected death rate for that portion of the population over that particular age. At successive ages the mean after-life in England and America for a half century ago is given in the following table:

[1] Forum, 26 : 540 (1899).

[2] New York City Health Dep't. Report, 1896, p. 14.

[3] Vital Statistics of Massachusetts, 1856–95, 812.

[4] Farr, Vital Statistics, 454. Eleventh Census, Vital Statistics, I : 484.

[5] N. A. Humphreys, in Royal Statistical Society *Journal*, 46 : 212 (1883).

Age.	Mean After Life (Males). England (Farr). 1838–54.	United States (Meech). 1830–60.
0	40	41
10	47	48
20	39	41
30	33	34
40	26	28
50	20	21
60	14	15
70	8	9
80	5	5

English and American mortality were thus nearly identical.

It will be noticed that the expectation of life does not always decline with advancing years. In infancy the reverse is the case, on account of the heavy mortality of that period. When the expectation does fall off, each year's advance of age reduces the expectation by less than a year. The advance in age and the reduction in expectation could only be equal if, within that year of age, there were no mortality whatever. A short rule for finding the expectation of life for any age between 20 and 45 is to expect to live beyond that age half way to age 96. For ages over 45, the expectation is half way to age 90.[1]

We have seen that the mean life-time has increased within the last half century. This has not been true of all the mean after-life-times. The effect of improved medical skill has been to prolong some weak lives, but only enough to shorten the average of survivors in general. Thus:

Age.	England.[2] 1838–54.	1881–90.
0	39.9	43.7
20	39.5	40.3
40	26.1	25.4
60	13.5	12.9
80	4.9	4.5

[1] Walford, Insurance Guide and Hand Book, 162.

[2] Vital Statistics of Massachusetts, 1856–95, 815.

Out of 100 aged 10 years, 66 used to survive to age 30; now 89 survive to that age. But 52 used to survive to age 60, and now only 51.[1]

In Massachusetts the death rates per different age-classes in 1865 and 1895 were:[2]

Ages	5–9	10–14	15–19	20–29	30–39	40–49	50–59	60–69	70–79	80–
1865	9.6	5.1	9.6	12.6	11.7	12.	17.	33	70	168
1895	6.2	3.2	5.3	7.1	9.7	13.	20.	39	82	185

It will be observed that for ages less than 40 the mortality for each age-class has fallen during the thirty years; but for higher ages, it has risen. It would be interesting to know the facts for urban and rural populations separately.

It has been claimed, in fact, that the effect of medical science is only to decrease infant mortality, to lengthen children's lives a little. But Humphreys has shown that for England two-thirds of the increased life is between the ages 20 and 60, that is during active life.[3]

A magnitude somewhat different from the expectation of life is the so-called probable after-life. It is not the most probable after-life, but that after-life which is as likely to be reached as not. To find it we take a survivorship table and mark that point in it at which just half of the original number are still living. Thus, if out of 1000 born, 500 will survive to age 45, the probable life for persons just born is 45. If, out of the 500 living at 45, 250 survive to age 70, the probable after-life for persons 45 years old is 25 years.

[1] Royal Statistical Society, *Journal*, 50: 649.

[2] Vital Statistics of Massachusetts, 1856–95, 755.

[3] Royal Statistical Society, *Journal*, 46: 213.

The expectation of life and the probable life for insured males in the United States are:[1]

Age.	Expectation.	Probable Life.
10	50	55
30	36	38
60	15	14

VI. LIFE TABLES.

Tables of survivorship or life tables are essential to the calculation of mean life times and true corrected death rates and, of course, to the work of the insurance actuary. They afford to the statistician the most powerful instrument he can have for the study of mortality, and should be regarded as the very kernel of mortality statistics.

The earliest attempt to discover a law for human vitality appears to be that of Ulpian, a Roman Praetorian Praefect, about 220 A.D. The meaning of his table is somewhat doubtful, but it is assumed to refer to expectation of life:[2]

Age.	Expectation of Life.
0–20	30
20–25	28
25–30	25
30–35	22
35–40	20
40–41	19[3]
49–50	10
50–55	9
55–60	7
60–	5

The Tuscan government is said to have authorized the use of Ulpian's table for valuing life annuities as late as 1814.[4]

[1] Meech, System and Tables of Life Insurance, p. *2.

[2] Walford, Insurance Guide and Handbook, 154.

[3] Between ages 40 and 50 the expectation decreases one year for each year of age. This could only be true, if there were no deaths between these ages. The table cannot, therefore, be accurate. (See last section, page 155).

[4] Assurance Magazine, 6: 314, note.

Graunt gave a table of survivorship based upon infererences from the burials in London between 1640 and 1660:[1]

Age.	Survivors.	Age.	Survivors.
0	100	46	10
6	64	56	6
16	40	66	3
26	25	76	1
36	16		

This was probably more conjecture than statistics.

Halley was the first to construct a life table of value in 1693, though Jean de Wit in Holland had, 22 years previously, provided a method of constructing such a table. Halley's table was based on the births and deaths of the city of Breslau, 1687–91, and deaths of the parish of Christ Church. After its publication there was great speculation in tontines. The first insurance company of importance, the Amicable, was started in 1706. Simpson continued Halley's work in 1742. Later Price constructed his Northampton and Carlisle tables. The former was based on deaths alone and the latter had too few data to be reliable. In 1766 Wargentin produced the first national table (Swedish). The best early tables are:[2]

Age.	Halley. 1687–91.	French; St. Maur. Before 1750.	Swedish; Wargentin. 1757–63.	Süssmilch. Before 1775.
0	1000	1000	1000	1000
10	661	484	611	532
20	598	449	570	491
40	445	314	459	374
60	242	168	293	210
80	41	23	56	37

Price constructed a table for Stockholm (1755–63) from which the following expectations of life were derived:[3]

[1] Walford, *Idem*, 154.

[2] Levasseur, in Royal Statistical Society *Journal*, 50 : 554.

[3] Walford, *Idem*, 155.

	Males.	Females.
0	14	18
5	31	37
10	30	37
20	24	30
40	16	19
60	9	10
75	4	4

The following tables are among the best for the last half of the 19th century. They relate to males:[1]

AGES.	0	5	10	20	40	60	80
Norway (Kiaer)	500	401	386	367	311	233	70
Denmark (David)	500	383	362	343	293	200	44
Sweden (Berg)	500	377	361	344	284	179	30
Switzerland (Gisi)	500	354	345	331	285	198	30
United States (Meech)	500	363	346	327	267	189	56
England (Farr)	500	361	345	326	265	178	40
France (Bertillon)	500	348	334	316	264	190	42
Belgium (Quetelet)	500	357	341	315	251	168	34
Prussia (Becker)	500	329	313	298	248	163	29
Italy (Bodio)	500	315	297	280	235	161	32
Bavaria (Hermann)	500	285	271	255	205	135	22

We note here important differences in the mortality of different nations, though a striking similarity in the main features. The United States and England are closely similar except at the later ages. The Scandinavian countries are in the lead, followed by those of Central Europe and America. Southern Europe (Italy) has low vitality.

The best general life table for the United States is probably that by Meech. It is based on the population returns for the censuses of 1830, 1840, 1850 and 1860, not on mortality statistics. It is not published in the census, having been prepared too late for the census of 1850, with which Mr. Meech was connected. Mr. Meech was not successful in inducing subsequent census super-

[1] M. Block, Traité de Statistique, 216; Bodio, Annualità Vitalizie, (Rome, 1892,) 2; Meech, System and Tables of Life Insurance, *258.

intendents to continue the calculations. The first census table was constructed by Elliott, government actuary, for the census of 1870. It can scarcely be accurate, being based on an assumed deficiency in the mortality record of 41 per cent. and being subject to much arbitrary adjustment. Its results, compared with those of Meech, are as follows:

	Expectation of Life.		
	Elliott.	Meech.	
Age.		Male.	Female.
0	39.	41.	43.
10	45.	48.	49.
40	26.	28.	29.

The Tenth Census was rich in life tables relating to the registration states and cities. Some were as follows:

	Expectation of Life. (White Males).			
	Massachusetts.	New Jersey.	Dist. of Columbia.	N. Y. city
0	44	46	41	33
10	51	52	47	45
40	29	29	26	24

About seventy sets of tables in all are presented in this census accompanied with interesting diagrams. Indeed the census was too prodigal as to quantity and somewhat careless as to quality. It is difficult to separate the wheat from the chaff. The tables should have been accompanied by a running criticism. A general defect was that no attempt was made to correct the deficiencies in the returns for infants. This correction is essential to a good life table.

The Eleventh Census devotes only three pages to life tables. One for New Jersey, based on six-years' records, gives expectations of life (for white males) as follows:

Ages.	Expectation.
0	40
10	49
40	27

The corresponding figures, from Meech, for the whole United States, 1830–60, are 41, 48, 28. The agreement

is closer than with the New Jersey table of 1880 (46, 52, 29), due perhaps to progressive precision in New Jersey registration. The population data were used as reported, though for Massachusetts and New Jersey additional tables were computed, in which the population under 15 years was adjusted by a formula given by Mr. Elliott, in the census of 1870. "Owing to the marked deficiency in the population data for children under 1 year of age . . . these life tables were even more inaccurate than those given in volume xii of the Tenth Census Reports, and the same must, therefore, be said of the expectation of life derived from them. The deficiency in the return of children under 1 year of age was so great that the omission assumed by Mr. Elliott as the basis of his formula in 1870 is entirely inadequate."[1]

Tables are given for the census year for Massachusetts, District of Columbia, Chicago, and Cincinnati, and for six years for New Jersey, Baltimore, Brooklyn, New York city, and Philadelphia.

Life tables are the basis of life insurance, but the tables employed by insurance companies to-day are constructed entirely from experience with insured lives. These are, of course, subject to medical selection. The expectations of life at age 40, for insured lives at different times, are as follows:[2]

French Tontines, Deparcieux, (1746)	27
French Tontines, Beauvisage, (1867)	29
French Ins. lives, Kertanguy, (1874)	27
French Pensioned Widows	29
Twenty English Companies, Males, (1869)	27
Twenty English Companies, Females, (1869)	28
English Gov't Pensions, Males	26
English Gov't Pensions, Females	30
German Gotha Company, (1880)	27
Thirty American Offices, Males	28

[1] Eleventh Census, Vital Statistics, I: 484.

[2] Levasseur, in Royal Statistical Society *Journal*, 50: 565 (1887).

These figures do not differ greatly from those derived from general mortality tables. The expectation of life at age 40 from the English Life Table Number 3, (1838–54), quoted above, is 26, and from the English table (1876–80) also 26; for the United States (1830–60) according to Meech, 28.

Selection among insured lives is not wholly in favor of the company. Those who elect to withdraw from a company are not a random collection, but consist largely of healthy lives. The unhealthy are almost sure to remain. Companies which charge a natural premium or assessment instead of a level one, are especially liable to this sort of adverse selection. The natural premium increases yearly. The level remains the same, so that as time goes on and the chances of death increase, the premium becomes progressively cheaper in proportion to the risk, and the insured has a stronger economic motive not to withdraw. For lack of precautions against withdrawals of their best risks, many companies have become bankrupt. In all companies, the mortality among lives recently insured is less than the tables allow, while among lives insured longer, the reverse is true. Thus, among 17 American offices the ratios of actual mortality to that computed from the tables were :[1]

For those insured one year or less	61 per cent
For those insured from five to six years	140 "

A corresponding study in England, made by Mr. E. Farren,[2] showed the following death rates per 1000:

Age.	General Population.	Insured Lives.	Insured only one year.
30	10.2	8.7	6.6
70	67.	63.	54.

[1] Emory McClintoch, (N. Y. Mutual Life), Effects of Selection, (New York, 1892), 31.

[2] Farr, Vital Statistics, 483.

Thus the influence of selection is in favor of the company at first and against it later on, while its general effect is to bring the mortality of insured lives only slightly below that of the general population.

The importance of selection in insurance is not yet appreciated by the companies. It belies the boasted precision of the actuary's calculation of risks. The time must come when its effect will have to be considered in calculating premiums. This will vastly complicate the actuary's work for, instead of relying on a single life table, he must have recourse to a separate life table for those insured at each age of life.

Some policies confine their holders to certain territories or latitudes. The wisdom of this restriction is seen by comparing the ordinary experience table with one constructed by Mr. C. N. Jones for American tropical experience :[1]

Age.	American Experience. 30 Offices.	Tropical Experience.
20	1000	1000
40	858	767
60	635	460
80	173	56

Statistics collected by Meikle, in England, and McClintock, in America, show greater longevity among abstainers from alcoholic beverages than among non-abstainers.

If the survivors opposite any age in a life table be deducted from the survivors of the year previous, the difference will be the numbers dying in the intervening year. A table of such differences is a mortality table. It is interchangeable with the survivorship table, since either can be obtained from the other. A complete set of life tables requires many columns besides the three thus far mentioned (survivors, dying, expectation of life).

[1] Meech, System and Tables, p. *2 ; Papers and Transactions Actuarial Society of America, No. 11, 1898, p. 316.

Sometimes as many as eleven columns are given and these do not include the "commutation tables" of the actuary, which involve the interest element.

The law of human mortality (or the equivalent law of survivorship) has usually been treated empirically. Exponential formulae have been fitted to it by Gompertz, Makeham, and others. Lexis, however, queried why it should not conform to the law of chance as, for instance, do the heights or weights of men. He observed that if the law of mortality be plotted graphically, the curve resembles, at the old age end, the curve of the law of chance. Recently Karl Pearson has shown that the entire mortality curve can be accounted for by adding together five chance curves.[1] The importance of this contribution to the subject lies in the fact that it rids it of its purely empirical character, and co-ordinates the law of human mortality with the law of the distribution of human heights, cranial capacity, and other measureable characteristics, as well as with the laws of lottery drawing, coin tossing, and games of chance. One curious point brought out by Pearson, is that the conformity of the mortality law with the law of chance is possible only by including ante-natal deaths. We must begin our observations at the beginning (conception), not after seven to ten months of life have passed.

The methods of constructing life tables can only be briefly mentioned here. It will not do simply to take statistics of, say 100,000 deaths, and record how many of the persons thus registered died 1 year old, 2 years old, etc. Such a table would not be a true table of survivorship, except in a stationary population. In an increasing population there is always an abnormal number of

[1] Chances of Death, and other Studies in Evolution, (London; Arnold), Vol. I, p. 26.

children and hence an abnormal number of children's deaths each year. Statistics of deaths in such a population would thus be overloaded with short lives. The mean duration of life calculated from such a table (*i. e.*, the "average age of the dying") would be too small. Dr. Price's Northampton table is an example of such an error. The mean duration of life was 25. Dr. Farr found for a later period the same mean life for Northampton, computed by the same (erroneous) method, but the true method gave 38![1]

If a population is stationary, however, statistics of death are sufficient to establish a life table. This method was used by Halley. As we have seen, statistics of age distribution in such a population would also give a life table. In practice, however, no stationary populations exist. The age distribution at successive censuses in a moving population (together with the total numbers) will suffice to determine a life table. This method is employed by Meech.

The best tables, however, are those constructed from the death rates for each age group. Knowing the death rate for the age group, 5–6, it is possible to compute how many out of a given number living at age 5 will survive to age 6. Beginning at age 0 (with any arbitrary number such as 100,000) the survivors can be recorded for each successive year to the end. The details of the method have been improved by numerous writers, including Fourier, Quetelet, Bertillon, Farr, Knapp, etc.

VII. SUGGESTIONS.

In closing this very rapid survey of the field of mortality statistics in the United States, the following recommendations for the twelfth census suggest themselves:

[1] Farr, Vital Statistics, 480.

1. Omit enumeration of deaths entirely. The resulting "statistics" are not only useless but mischievous. They cast discredit on the census as a whole. They increase its cost and occupy valuable space. Their collection and analysis absorb attention which would otherwise be free to prepare the really valuable features of the census.

2. Accompany every figure subject to doubt with a statement as to its possible or probable error, such notes to be attached to the tables and figures by reference stars or letters.

3. Dispense with all digits not really significant.

4. Calculate rates on basis of population at center of year.

5. Make more use of corrected death rates.

6. Make more use of life tables, but only where they can be accurately computed.

7. Give prominence to age distribution tables in connection with life tables, using solid diagrams such as Bodio's.[1]

8. Make the census of 1900 a Jubilee census and review critically all the mortality statistics of value (not great in amount) in the United States, official and unofficial, with a briefer historical review of the best work abroad.

9. Insert short catch titles for the tables, and in general make tables of contents and indexes more usable.

A more radical recommendation is to establish a national system of registration of births, marriages and deaths. This need not pretend at first, to be complete but only representative.[2] Certain areas may be selected for national registration in the south, west, etc., and

[1] Annali di Statistica, 12 : 16,ff., (1880).

[2] C. L. Wilbur, in Am. Stat. Assn. *Publications*, 5 : 188 (1897).

the results compared with state registrations in the east. As Dr. Wilbur says, in meteorology we establish selected stations of observation. We do not attempt to record all the weather. Attempts at completeness are apt to ruin accuracy. There will almost certainly be some bad figures to contaminate the good. An essential feature of Dr. Wilbur's plan would be to empower the registrar to reject returns shown to be seriously defective.

VIII. LITERATURE.

The following list is by no means a complete bibliography. To enumerate all that has been written on mortality statistics (even excluding their application to life insurance) would require many hundred pages. The writings noted below are merely a few of such as would be most useful to the American student and would put him on the track of earlier and collateral references. Among important writers omitted are De Wit, Halley, Süssmilch, De Parcieux, Price, Tetens, De Moivre, Fourier, La Place, Baily, Milne, Moser, De Morgan, Becker, Wittstein, Gompertz, Makeham, Bodio, Körösi, Billings.

United States Census Volumes, beginning with 1850. Only those of 1880 and 1890 are of much value.

Annual reports of the National, state, and municipal boards of health, especially those of Massachusetts.

A Summary of the Vital Statistics of the New England States for the year 1892, (Boston, Damrell & Upham), by the secretaries of the state boards of health.

The Vital Statistics of Massachusetts, 1856–95 (in Twenty-Eighth Annual Report of State Board of Health of Massachusetts for 1896), by S. W. Abbott, Secretary.

Bills of Mortality, 1810–1849, City of Boston, by Lemuel Shattuck, Boston (Registry Department), 1893, xliv, 87 pp.

Monthly Bulletin of Vital Statistics of Connecticut, New York, Michigan, and other states.

Publications of the American Public Health Association.

Publications of the American Statistical Association, especially 1892–3.

Race Traits and Tendencies of the American Negro, by F. L. Hoffman. In Publications of the American Economic Association, 1896.

Publications of the American Society of Actuaries.

The Probabilities of the Duration of Human Life in the United States, Am. Philos. Soc., Philadelphia, Vol. III, p. 25, 1st Series, 1791. Wm. Barton.

Philadelphia Life Tables, Am. Philos. Soc. of Philadelphia, 1870, p. 17. By P. E. Chase.

Social Science Association of Philadelphia, 1879. Stockton-Hough on "Statistics of Births, Deaths, Marriages, etc., in Philadelphia, 1861–72," and "Relative Influence of City and Country Life," etc.

System and Tables of Life Insurance, by Levi W. Meech.

Statistics and Sociology, by Richmond Mayo-Smith, New York, 1895.

Reports of the Registrar of Vital Statistics of England and other countries.

Journal of the Royal Statistical Society: *e.g.*, "Value of Death Rates," December, 1874, by Humphreys; "Decline in English Death Rate," 1883, by Humphreys; "Tables of Mortality and Survivorship," by Levasseur, March, 1887; "Summary of Several Male Life Tables," 1887, by Ogle.

Journal of the Institute of Actuaries, *e.g.*, "Contributions to the History of Life Insurance," 1850; "Construction of Life Tables," 1861, by Farr.

Text-book of the Institute of Actuaries.

Insurance Guide and Handbook, by C. Walford, 2d Edition, 1867.

New Life Table of Glasgow, by H. Chalmers.

English Life Table, by Wm. Farr. (This is the best life table extant).

Vital Statistics, by Wm. Farr, London, 1885.

Elements of Vital Statistics, by A. Newsholme, London, 1889.

Studies in Statistics, by G. B. Longstaff, London, 1891.

The Chances of Death, etc., by Karl Pearson, London, 1897.

Bulletin de l'Institut International de Statistique.

Journal de la Société de Statistique de Paris, *e.g.*, "Diverses manières de mesurer la vie humaine," March, 1886, by Bertillon.

Allgemeines statistisches Archiv.

Annales de Demographie Internationale.

Annualita vitalizie, Rome.

Annali di Statistica, Rome.
Ueber die Ermittlung der Sterblichkeit, etc., 1868, by G. F. Knapp.
Abhandlungen aus der mathematischen Statistik, Leipzig, 1869, by G. Zeuner.
Physique sociale, by Ad. Quetelet.
Tables de mortalité et leur developpement, 1872, by Quetelet.
Einleitung in die Theorie der Bevölkerungsstatistik, Strassburg, 1875, by W. Lexis.
Traité theorique et pratique de statistique, Paris, 1878, by M. Block.
Die Lehre von der Mortalität und Morbilität, Jena, 1882, by H. Westergaard.
Die mittlere Lebensdauer, Jena, 1893, by L. von Bortkewitsch.
Statistik und Gesellschaftslehre, 1895–7, by G. von Mayr.

IRVING FISHER.

Yale University.